The effect of assuming the identity as a generator on the length of the matrix algebra

Thomas Laffey[a], Olga Markova[b,2], Helena Šmigoc[a,*]

[a] *School of Mathematical Sciences, University College Dublin, Belfield, Dublin 4, Ireland*
[b]*Department of Algebra, Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, Moscow 119991, Russia*

Abstract

Let $M_n(\mathbb{F})$ be the algebra of $n \times n$ matrices and let $\mathcal{S}$ be a generating set of $M_n(\mathbb{F})$ as an $\mathbb{F}$-algebra. The length of a finite generating set $\mathcal{S}$ of $M_n(\mathbb{F})$ is the smallest number k such that words of length not greater than k generate $M_n(\mathbb{F})$ as a vector space. Traditionally the identity matrix is assumed to be automatically included in all generating sets $\mathcal{S}$ and counted as a word of length 0. In this paper we discuss how the problem changes if this assumption is removed.

Keywords: Finitely generated algebra, Lengths of Algebras
2010 MSC: 15A30, 13E10, 16S50.

This paper is dedicated to the memory of Hans Schneider.

1. Introduction

Let $\mathcal{A}$ be a finite-dimensional $\mathbb{F}$-algebra for some field $\mathbb{F}$. The *length of a finite generating set $\mathcal{S}$* of algebra $\mathcal{A}$ is the smallest number k such that words of length not greater than k generate $\mathcal{A}$ as a vector space. The *length of the algebra* is the maximum of lengths of its generating sets. The problem of evaluating the length of the full matrix algebra was posed by A. Paz in [13], where it is conjectured that the length of the full matrix algebra $M_n(\mathbb{F})$

---

*Corresponding author

Email addresses: `thomas.laffey@ucd.ie` (Thomas Laffey), `ov_markova@mail.ru` (Olga Markova), `helena.smigoc@ucd.ie` (Helena Šmigoc)

is equal to $2n - 2$. While several results on the length of $M_n(\mathbb{F})$ [12, 8, 6], its generating sets [7, 1, 9] and subalgebras [10, 3, 11, 4] were found, the conjecture remains open.

It is usually assumed, that the identity matrix is included in generating set $\mathcal{S}$ and is regarded as having length 0. In this paper we pose the question, how do the lengths of generating sets and of algebras differ if we remove this assumption. While we can talk about this notion for more general algebras, we will concentrate our discussion on matrix algebras.

Let us begin by introducing some notation. Let $\mathcal{S}$ be a subset of an algebra $\mathcal{A}$, then $\mathcal{L}(\mathcal{S})$ will denote the subalgebra of $\mathcal{A}$ generated by the set $\mathcal{S}$, and $\mathcal{L}_k(\mathcal{S})$ will denote the vector space generated by all the words in $\mathcal{S}$ of length at most k. If $\mathcal{A}$ is unitary then $\mathcal{L}_0(\mathcal{S}) = \mathbb{F}$ and $\mathcal{L}_0(\mathcal{S}) = 0$ otherwise. Clearly we have:

$$\mathcal{L}(\mathcal{S}) = \bigcup_{i=1}^{\infty} \mathcal{L}_i(\mathcal{S}).$$

The length of a generating set $\mathcal{S}$ will be denoted by $l(\mathcal{S})$, and the length of the algebra $\mathcal{A}$ will be denoted $l(\mathcal{A})$. Note that $l(\mathcal{S})$ is the smallest number k, such that $\mathcal{L}_k(\mathcal{S}) = \mathcal{L}_{k+1}(\mathcal{S})$.

Now we introduce the same notions as above again, but this time we do not assume that the identity is included in the generating set $\mathcal{S}$. For a subset $\mathcal{S}$ of an algebra $\mathcal{A}$ we will denote $\mathcal{L}_k^0(\mathcal{S})$ to be the vector space generated by all the words in $\mathcal{S}$ of length at most k, where we do not assume the identity is included in $\mathcal{S}$. (So in general $\mathcal{L}_k^0(\mathcal{S}) \subseteq \mathcal{L}_k(\mathcal{S})$.) Similarly, $\mathcal{L}^0(\mathcal{S})$ will denote the algebra generated by the set $\mathcal{S}$. We also define the notion of the length in this sense. The length of a generating set $\mathcal{S}$ without the identity automatically included in $\mathcal{S}$ will be denoted by $l_0(\mathcal{S})$ and $l_0(\mathcal{A})$ will denote the length of $\mathcal{A}$ in this sense.

We will also need some notation for matrices. $J_k(\lambda)$ will denote the $k \times k$ Jordan block corresponding to the eigenvalue λ, and $J_k = J_k(0)$. We will denote by I_n the $n \times n$ identity matrix, by O_n, $O_{r,s}$ the $n \times n$ and $r \times s$ zero matrices and by $E_{i,j}$ the matrix unit, i.e. the matrix with 1 in (i, j)-th position and 0 elsewhere.

2. Main Results

The following proposition is perhaps the main motivation, why the identity is traditionally assumed to be included the generating set.

Proposition 2.1. *Let $\mathcal{S}$ be a generating set of a simple algebra $\mathcal{A}$ of dimension greater than 1, i.e. $\mathcal{L}(\mathcal{S}) = \mathcal{A}$. Then $\mathcal{L}^0(\mathcal{S}) = \mathcal{A}$.*

Proof. Since $\mathcal{L}^0(\mathcal{S})$ is a nonzero ideal in $\mathcal{A}$ the statement follows. □

Let $\mathcal{S}$ be a generating set of $M_n(\mathbb{F})$. Then, clearly $l_0(\mathcal{S}) = l(\mathcal{S})$ or $l_0(\mathcal{S}) = l(\mathcal{S}) + 1$. Simple examples below show that both eventualities can occur.

Example 2.2. Let us consider a generating set

$$\mathcal{S} = \{E_{12}, E_{21}, E_{22}\} \subset M_2(\mathbb{F}).$$

Then $\mathcal{L}(\mathcal{S}) = M_2(\mathbb{F})$, $l(\mathcal{S}) = 1$ and $l_0(\mathcal{S}) = 2$. For $\mathcal{S}' = \mathcal{S} \cup \{E_{11}\}$ we have $l(\mathcal{S}) = 1$ and $l_0(\mathcal{S}) = 1$.

Next set of results preset some cases, for which $l(\mathcal{S}) = l_0(\mathcal{S})$.

Proposition 2.3. *Let $\mathcal{S}$ be a generating set for $M_n(\mathbb{C})$.*

1. *If $l(\mathcal{S}) \geq n$ and if $\mathcal{L}_1^0(\mathcal{S})$ contains an invertible matrix, then $l(\mathcal{S}) = l_0(\mathcal{S})$.*

2. *If $l(\mathcal{S}) = 2n - 2$ and if $\mathcal{L}_2^0(\mathcal{S})$ contains a derogatory invertible matrix, then $l(\mathcal{S}) = l_0(\mathcal{S})$.*

Proof. Let $A \in M_n(\mathbb{C})$ be an invertible matrix in $\mathcal{L}_1^0(\mathcal{S})$. By the Cayley-Hamilton theorem: $A^n + a_1 A^{n-1} + \ldots + a_n I_n = 0$, where for $i = 1, \ldots, n$: $a_i \in \mathbb{C}$ and $a_n \neq 0$. Therefore:

$$I_n = -\frac{1}{a_n} A^n - \frac{a_1}{a_n} A^{n-1} - \ldots - \frac{a_{n-1}}{a_n} A$$

is contained in $\mathcal{L}_n^0(\mathcal{S})$.

To prove the second item, let A be a derogatory invertible matrix in $\mathcal{L}_2^0(\mathcal{S})$. Its minimal polynomial has degree at most $n - 1$ so:

$$A^k + a_1 A^{k-1} + \ldots + a_k I_n = 0$$

for $k \leq n - 1$ and $a_k \neq 0$. Now

$$I_n = -\frac{1}{a_k} A^k - \frac{a_1}{a_k} A^{k-1} - \ldots - \frac{a_{k-1}}{a_k} A$$

is contained in $\mathcal{L}_{2n-2}^0(\mathcal{S})$. □

The two lengths are the same for the matrix algebras $M_2(\mathbb{F})$ and $M_3(\mathbb{F})$. The equations $l(M_2(\mathbb{F})) = 2$ and $l(M_3(\mathbb{F})) = 4$ follow from [13].

Theorem 2.4. $l_0(M_2(\mathbb{F})) = 2$.

Proof. Let $\mathcal{S}$ be a generating set for $M_2(\mathbb{F})$ that does not contain I_2. Trivially we can replace it by an irredundant subset which generates $M_2(\mathbb{F})$, so we can assume $\mathcal{S}$ is irredundant. Then $\mathcal{S}$ has either 2 or 3 elements. If $\mathcal{S}$ has three elements, then $\mathcal{S} \cup \{I_2\}$ is a spanning set for $M_2(\mathbb{F})$ and since then $\mathcal{L}_1(\mathcal{S}) = M_2(\mathbb{F})$, the result holds.

Now suppose that $\mathcal{S}$ has two elements, say $\mathcal{S} = \{A, B\}$. If $\mathcal{L}_1^0(\mathcal{S})$ contains an invertible matrix, then the result is true by Proposition 2.3. So, arguing by contradiction, we may assume $A^2 = \alpha A$ and $B^2 = \beta B$, for some $\alpha, \beta \in \mathbb{F}$. If $\alpha = 0$, then we may without loss of generality assume that

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Now let

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Since $\mathcal{L}_1^0(\mathcal{S})$ does not contain an invertible matrix, we have $\det(xA + B) = 0$ for every $x \in \mathbb{F}$, and this implies, that $b_{21} = 0$. This contradicts the fact that $\mathcal{S}$ generates $M_2(\mathbb{F})$. So suppose that $\alpha \neq 0$. Replacing A by $\frac{1}{\alpha}A$ we may assume that $A^2 = A$ and using similarity we may assume that

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Now $\det(xA + B) = 0$ for all $x \in \mathbb{F}$ implies $b_{22} = 0$ and since $\mathcal{S}$ generates $M_2(\mathbb{F})$, we must have $b_{21}b_{12} \neq 0$. But then $\det B \neq 0$ contradicting our earlier assumption. $\qquad\square$

Theorem 2.5. $l_0(M_3(\mathbb{F})) = 4$.

Proof. It is sufficient to prove that $l_0(\mathcal{S}) \leq 4$ for any generating set $\mathcal{S}$. If $\mathcal{S}$ contains an invertible matrix, then the result follows from Proposition 2.3.

Next we assume that $\mathcal{S}$ contains a singular nonderogatory matrix A. We will distinguish several cases depending on the size of the Jordan block associated with 0. First consider the case when the Jordan canonical form of A is J_3 or A has a singular Jordan block of size 2 and an eigenvalue $a \in \mathbb{F} \setminus \{0\}$. In this case we may without loss of generality assume that

$A = J_3$ or $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a \end{pmatrix}$. First consider the sequence of inclusions:

$$\mathbb{F}A^2 \subseteq \mathcal{L}_1(\mathcal{S})A^2 \subseteq \ldots \subseteq \mathcal{L}(\mathcal{S})A^2 = M_3(\mathbb{F})A^2.$$

We have $A^2 = E_{13}$ or $A^2 = a^2 E_{33}$, i.e. A^2 is a rank one matrix, thus this sequence stabilises at most on the step 2, so we get: $\mathcal{L}_2(\mathcal{S})A^2 = M_3(\mathbb{F})A^2$. Furthermore, $\mathcal{L}_2(\mathcal{S})A^2 \subseteq \mathcal{L}_4^0(\mathcal{S})$, hence we have proved that all matrices with the first two columns equal to zero are contained in $\mathcal{L}_4^0(\mathcal{S})$. We repeat the inclusion argument for A, yielding that $\mathcal{L}_4^0(\mathcal{S})$ contains all matrices with the first column equal to zero. Now we choose $X = (x_{ij}) \in \mathcal{S}$ with $x_{31} \neq 0$ or (if such $X \in \mathcal{S}$ does not exists), we choose $X = (x_{ij}) \in \mathcal{S}$ with $x_{21} \neq 0$ and $x_{31} = 0$. Since $\mathcal{S}$ is a generating set for $M_3(\mathbb{F})$ such a matrix X exists. Then $\mathcal{L}_2(\mathcal{S})X$ projected to column 1 is $\mathbb{F}^3$ and since $\mathcal{L}_2(\mathcal{S})X \subseteq \mathcal{L}_3^0(\mathcal{S})$, we conclude $M_3(\mathbb{F}) \subseteq \mathcal{L}_4^0(\mathcal{S})$ as asserted.

If A has a singular Jordan block of order 1, then we may assume that it is of the form:

$$A = \begin{pmatrix} 0 & 0 \\ 0 & B \end{pmatrix},$$

where B is a 2×2 nonsingular matrix. It follows from the proof of [11, Proposition 3.19] that it would not affect the length of a given generating set if we take an extension of $\mathbb{F}$ as the base field. Therefore we may assume that B is either equal to $J_2(a)$ or to a diagonal matrix with two distinct elements a and b. In the first case we notice that

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \frac{1}{a}B^2 - B \in \mathcal{L}_2^0(\{B\})$$

and that

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = -\frac{1}{a^2}B^2 + \frac{2}{a}B \in \mathcal{L}_2^0(\{B\}).$$

In the second case we have

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \frac{a}{b(a-b)}B - B^2 \in \mathcal{L}_2^0(\{B\})$$

and

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{b}{a(b-a)}B - B^2 \in \mathcal{L}_2^0(\{B\}).$$

In both cases we may use the inclusion argument to prove that $\mathcal{L}_4^0(\mathcal{S})$ contains all matrices with the first column equal to zero. As we did before, choosing an $X \in \mathcal{S}$ with x_{j1}, $j \in \{2, 3\}$, completes the proof.

From now on we may assume that all the elements in $\mathcal{L}_1^0(\mathcal{S})$ are singular and derogatory. This implies that for all $A \in \mathcal{S}$ we have $A^2 = \alpha(A)A$, for some $\alpha(A) \in \mathbb{F}$. In [5, 2] it is proved, that in this case $\mathcal{S}$ has to contain at least three elements in order to generate $M_3(\mathbb{F})$.

First let us assume that there exists $A \in \mathcal{S}$ with $\alpha(A) = 0$. Without loss of generality we may assume that:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Let B be another element in $\mathcal{S}$. Then:

$$(A + B)^2 = \alpha(A + B)(A + B),$$

and this implies that

$$AB + BA = \alpha(A + B)A + (\alpha(A + B) - \alpha(B))B.$$

Taking $B = (b_{ij})$ and using the form of A, this gives us:

$$\begin{pmatrix} b_{21} & b_{22} + b_{11} & b_{23} \\ 0 & b_{21} & 0 \\ 0 & b_{31} & 0 \end{pmatrix} = \beta \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \gamma \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix},$$

where $\beta = \alpha(A + B)$ and $\gamma = (\alpha(A + B) - \alpha(B))$. Now, if $\gamma \neq 0$, then $b_{21} = b_{31} = 0$, and the same is true if $\gamma = 0$. This proves, that all matrices in $\mathcal{S}$ are of the form:

$$\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ 0 & b_{22} & b_{23} \\ 0 & b_{32} & b_{33} \end{pmatrix},$$

and this contradicts our assumption that $\mathcal{S}$ generates $M_3(\mathbb{F})$.

Now we can assume that $\alpha(A) \neq 0$ for all $A \in \mathcal{S}$. First consider the case when there exists a matrix of rank 1 in $\mathcal{L}_1^0(\mathcal{S})$. Without loss of generality we may assume that there exists

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathcal{S}.$$

Let B be another matrix in $\mathcal{S}$. We again look at $(A+B)^2 = \alpha(A+B)(A+B)$:

$$\begin{pmatrix} 2b_{11} & b_{12} & b_{13} \\ b_{21} & 0 & 0 \\ b_{31} & 0 & 0 \end{pmatrix} = \beta \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \gamma \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix},$$

where $\beta = \alpha(A + B) - 1$ and $\gamma = (\alpha(A + B) - \alpha(B))$. If $\gamma = 0$, then the above formula implies $b_{12} = 0$, $b_{13} = 0$, $b_{21} = 0$ and $b_{31} = 0$. If $\gamma \neq 0$, then we have $b_{22} = 0$, $b_{23} = 0$, $b_{32} = 0$ and $b_{33} = 0$. In this case $B^2 = \alpha(B)B$, gives us $b_{21}b_{12} = 0$, $b_{21}b_{13} = 0$, $b_{31}b_{12} = 0$ and $b_{31}b_{13} = 0$. Now, if $b_{21} \neq 0$ or if $b_{31} \neq 0$, then $b_{12} = 0$ and $b_{13} = 0$, otherwise $b_{21} = 0$ and $b_{31} = 0$. With this we have proved that all the matrices in $\mathcal{L}_1^0(\mathcal{S})$ have one of the following forms:

$$\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ 0 & b_{22} & b_{23} \\ 0 & b_{32} & b_{33} \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}. \tag{1}$$

But since for any two matrices B and C in $\mathcal{S}$ matrices B, C and $B+C$ have to have one of the forms above, they all have to have the same form. This contradicts our assumption that $\mathcal{S}$ generates $M_3(\mathbb{F})$.

Now consider the case that all matrices in $\mathcal{L}_1^0(\mathcal{S})$ have rank 2. Without loss of generality we may assume that there exists $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathcal{S}$.

Then $I_3 - A = E_{11}$. Since all matrices from $\mathcal{L}_1^0(\mathcal{S})$ are derogatory, then all the matrices from the space $\mathcal{L}_1(\mathcal{S})$ share this property, therefore all the matrices from the space $\mathcal{L}_1^0((\mathcal{S} \setminus \{A\}) \cup \{I_3 - A\})$ also are derogatory.

Let B be another matrix in $\mathcal{S}$. As above we consider $((I_3 - A) + B)^2 = \alpha(I_3 - A + B)(I_3 - A + B) + \delta I_3$:

$$\begin{pmatrix} 2b_{11} & b_{12} & b_{13} \\ b_{21} & 0 & 0 \\ b_{31} & 0 & 0 \end{pmatrix} = \beta \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \gamma \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} + \delta \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where $\beta = \alpha(I_3 - A + B) - 1$, $\gamma = (\alpha(I_3 - A + B) - \alpha(B))$. If $\gamma = 0$, then the above formula implies $b_{12} = 0$, $b_{13} = 0$, $b_{21} = 0$ and $b_{31} = 0$. If $\gamma \neq 0$, then we have $b_{23} = 0$ and $b_{32} = 0$, $b_{22} = b_{33}$. If $b_{13} \neq 0$, in this case $B^2 = \alpha(B)B$, gives us $(B^2)_{23} = b_{21}b_{13} = 0$, i.e. $b_{21} = 0$, $(B^2)_{22} = (B^2)_{33}$, i.e. $b_{31} = 0$. Analogously, if $b_{31} \neq 0$, then $B^2 = \alpha(B)B$ implies that $b_{12} = b_{13} = 0$. If both $b_{13} = b_{31} = 0$, then $B^2 = \alpha(B)B$ gives us that $(B^2)_{22} = b_{12}b_{21} + b_{22}^2 = b_{22}^2 = (B^2)_{33}$, thus $b_{12}b_{21} = 0$. Again, we have proved that all matrices in $\mathcal{S}$ have block forms as in (1), a contradiction.

□

Next we show that even generating sets $\mathcal{S}$ of $M_n(\mathbb{F})$ with relatively large lengths can have $l_0(\mathcal{S}) = l(\mathcal{S}) + 1$.

Definition 2.6. For $k = -(n-1), -(n-2), \ldots, n-2, n-1$ we will say the the elements of the form a_{ii+k} form the $k-diagonal$ of a matrix $A \in M_n(\mathbb{F})$.

Lemma 2.7. *Let A be a matrix with all its nonzero elements on its k-diagonal, and let B be a matrix with all its nonzero elements on its l-diagonal. Then AB can have nonzero entries (if any) only on its $k + l$-diagonal.*

Theorem 2.8. *Let $B_n = E_{n-1,1} - E_{n,2} \in M_n(\mathbb{C})$, $n \geq 3$. Then*

$$\mathcal{L}(\{J_n, B_n\}) = M_n(\mathbb{C}),$$

$l(\{J_n, B_n\}) = 2n - 3$ *and* $l_0(\{J_n, B_n\}) = 2n - 2$.

Proof. Since $B_n^2 = 0$, every word in J_n and B_n is of the form:

$$B_n^{i_1} J_n^{j_1} B_n J_n^{j_2} B^{i_2} \ldots B_n J_n^{j_s} B_n^{i_2}, \tag{2}$$

where $i_1, i_2 \in \{0, 1\}$, and j_l, $l = 1, 2, \ldots, s$, are positive integers. Lemma 2.7 tells us that a word of the form (2) can have nonzero entries (if any) only on $\sum_{l=1}^{s} j_s - (n-2)(i_1 + i_2 + s - 1)$-diagonal.

First we show that $l_0(\{J_n, B_n\}) > 2n - 3$. To show this, we consider the 0-diagonal of matrices in $\mathcal{L}_{2n-3}^0(\{J_n, B_n\})$. For a word (2) to have entries on the 0-diagonal, we need to have:

$$\sum_{l=1}^{s} j_s - (n-2)(i_1 + i_2 + s - 1) = 0,$$

hence at least one B_n has to appear in the product. All the products with nonzero entries on the 0-diagonal involving only one B_n are of the form:

$$J_n^j B_n J_n^{n-2-j} \text{ for } j = 0, 1, \ldots, n - 2.$$

There are $n - 1$ such products, and since the 0-diagonal has dimension n, we need at least one word involving 2 matrices B_n, in order to have the complete 0-diagonal included in the span. For this word to have elements on 0-diagonal we need $\sum_{l=1}^{s} j_s - (n-2)(2) = 0$, hence $\sum_{l=1}^{s} j_s = 2n - 4$. The length of this word is $2n - 2$, and this proves $l_0(\{J_n, B_n\}) > 2n - 3$.

Now we show that $l(\{J_n, B_n\}) = 2n-3$, which also implies that $l_0(\{J_n, B_n\}) = 2n - 2$.

Let $\mathcal{V}_n(k)$, $k = -(n-1), -(n-2), \ldots, (n-2), (n-1)$, denote the vector space of $n \times n$ matrices with nonzero entries on k-diagonal. Note that $\dim(\mathcal{V}_n(-(n-k))) = \dim(\mathcal{V}_n(n-k)) = k$ for $k = 1, \ldots, n$. For each of the vector sets $\mathcal{V}_n(k)$ we will find a basis $\mathcal{B}_n(k)$ consisting of words in J_n and B_n of length less than or equal to $2n - 3$.

First we compute $B_n J_n^{n-3} B_n = -E_{n,1}$, and this tells us that

$$\mathcal{B}_n(-(n-1)) = \{B_n J_n^{n-3} B_n\}$$

is a basis for $\mathcal{V}_n(-(n-1))$. Next we compute:

$$B_n J_n^{n-3} B_n J_n^{k-1} = -E_{n,k}$$

for $k = 2, \ldots, n-1$ and for $s \le n-2$, $t \le n-2$:

$$J_n^s B_n = E_{n-s-1,1} - E_{n-s,2} \text{ and } J_n^s B_n J_n^t = E_{n-s-1,t+1} - E_{n-s,t+2}.$$

Now it is clear that

$$\begin{aligned}
\mathcal{B}_n(-(n-k)) &= \{B_n J_n^{n-3} B_n J_n^{k-1}\} \cup \{J_n^l B_n J_n^{k-2-l};\ l = 0, \ldots, k-2\} \\
&= \{-E_{n,k}\} \cup \{E_{n-l-1,k-l-1} - E_{n-l,k-l};\ l = 0, \ldots, k-2\}
\end{aligned}$$

is a basis for $\mathcal{V}_n(-(n-k))$ for $k = 2, \ldots, n-1$.

To form $\mathcal{B}_n(0) = \mathcal{B}_n(n-n)$ we replace the word $B_n J_n^{n-3} B_n J_n^{n-1}$ of length $2n-2$ with I_n:

$$\begin{aligned}
\mathcal{B}_n(0) &= \{I_n\} \cup \{J_n^l B_n J_n^{n-2-l};\ l = 0, \ldots, n-2\} \\
&= \{I_n\} \cup \{E_{n-l-1,n-l-1} - E_{n-l,n-l};\ l = 0, \ldots, n-2\}
\end{aligned}$$

Note, that the set $\mathcal{B}_n(0)$ is linearly independent over any field of characteristics 0.

To get a basis for $\mathcal{V}_n(n-k)$ we also need to compute:

$$J_n^{n-1} B J_n^s = -E_{1,s+2}$$

for $s = 0, \ldots, n-3$. Now it is clear that

$$\begin{aligned}
\mathcal{B}_n(n-k) &= \{J_n^l B_n J_n^{2n-k-2-l}; l = n-k, \ldots, n-1\} \\
&= \{E_{n-l-1,2n-k-l-1} - E_{n-l,2n-k-l}; l = n-k, \ldots, n-2\} \cup \{-E_{1,n-k+1}\} = \\
&= \{E_{k-1-j,n-j-1} - E_{k-j,n-j}; j = 0, \ldots, k-2\} \cup \{-E_{1,n-k+1}\}
\end{aligned}$$

is a basis for $\mathcal{V}_n(n-k)$ for $k = 2, \ldots, n-1$. Finally we take

$$\mathcal{B}_n(n-1) = \{J_n^{n-1}\}.$$

Since all the words in $\mathcal{B}_n(l)$, $l = -(n-1), -(n-2), \ldots, (n-2), (n-1)$ have length less than or equal to $2n-3$, we have proved our claim. $\square$

We conclude this paper by exhibiting a class of generating sets for $M_n(\mathbb{C})$ with two elements, akin to the one in Theorem 2.8.

Theorem 2.9. *Let $i < n$ be positive integers. $\mathcal{L}(\{J_n^i, (J_n^T)^{n-i}\}) = M_n(\mathbb{C})$ if and only if $n-i$ and i are co-prime. If $\gcd(n, i) = d$, then $\mathcal{L}(\{J_n^i, (J_n^T)^{n-i}\}) \cong M_{\frac{n}{d}}(\mathbb{C})$.*

Proof. We will show that, for i and $n-i$ co-prime, $\mathcal{L}(\{J_n^i, (J_n^T)^{n-i}\})$ has no invariant subspaces except $\{0\}$ and $\mathbb{C}^n$, thereby showing that $\mathcal{L}(\{J_n^i, (J_n^T)^{n-i}\}) = M_n(\mathbb{C})$ by Burnside's Theorem.

Let
$$v = \begin{pmatrix} a_1 & a_2 & \dots & a_k & 0 & \dots & 0 \end{pmatrix}^T \in \mathbb{C}^n \qquad (3)$$

with $a_k \neq 0$. If $k \leq i$, then $(J_n^T)^{n-i}v$ has a nonzero element in $(k+n-i)$-th position, and if $k \geq i+1$, then $J_n^i v$ has a nonzero element in $(k-i)$-th position. We consider the following recursively defined sequence with $k_0 = k$ and:
$$k_{t+1} = \begin{cases} k_t + (n-i) & \text{for } k_t \leq i \\ k_t - i & \text{for } k_t \geq 1+i \end{cases} . \qquad (4)$$

Let j_0 be the smallest integer at which the sequence repeats: $k_{i_0} = k_{j_0}$ for some $i_0 < j_0$. Then: $k_{j_0} = k_{i_0} - \alpha i + \beta(n-i)$ and $\alpha i = \beta(n-i)$. Since i and $n-i$ are co-prime, we need $\alpha = \gamma(n-i)$ and $\beta = \gamma i$. This tells us that the sequence repeats after exactly n steps, independent of where we start, and it contains all numbers $\{1, 2, \dots, n\}$ is some order.

Let $\mathcal{U}$ be an invariant subspace for $\mathcal{L}(\{J_n^i, (J_n^T)^{n-i}\})$, and let v be a vector of the form (3) in $\mathcal{U}$ with $a_k \neq 0$. Since $\mathcal{U}$ has to contain an eigenvector of J_n^i we know that such a vector exists in $\mathcal{U}$ with $k \leq i$. Let k_i be the sequence defined in (4). Let $v_0 = v$ and let

$$v_{t+1} = \begin{cases} (J_n^T)^{(n-i)} v_t & \text{if } k_t \leq i \\ (J_n)^i v_t & \text{if } k_t \geq 1+i \end{cases} . \qquad (5)$$

Notice that v_i has a nonzero element a_k in the k_i-th position followed by zero elements. Taking into account our argument about the sequence k_i above, we conclude that $\{v_i; i = 0, \dots, n-1\}$ forms a basis for $\mathbb{C}^n$, thus proving $\mathcal{U} = \mathbb{C}^n$.

Now assume that $\gcd(i, n-i) = d$, $d \neq 1$. We can consider matrices $J_n^i, (J_n^T)^{n-i}$ as block $\frac{n}{d} \times \frac{n}{d}$ matrices with blocks from $M_d(\mathbb{C})$. We have all blocks in J_n^i and $(J_n^T)^{n-i}$ equal to either O_d or I_d, hence

$$J_n^i = J_{\frac{n}{d}}^{\frac{i}{d}} \otimes I_d, \ (J_n^T)^{n-i} = (J_{\frac{n}{d}}^T)^{\frac{n-i}{d}} \otimes I_d.$$

Therefore the algebra $\mathcal{L}(\{J_n^i, (J_n^T)^{n-i}\})$ is naturally isomorphic to $\mathcal{L}(\{J_{\frac{n}{d}}^{\frac{i}{d}}, (J_{\frac{n}{d}}^T)^{\frac{n-i}{d}}\})$ which is $M_{\frac{n}{d}}(\mathbb{C})$ as shown above. $\square$

Lemma 2.10. *Let n, i, be positive integers, $n \geq 2$, $i < n$, $\gcd(n, i) = 1$. Let $J_n^i, (J_n^T)^{n-i} \in M_n(\mathbb{C})$. Denote*

$$\mathcal{W}(n, i) = \begin{cases} \mathcal{V}(n - 2i) \oplus \mathcal{V}(-2i), & \text{if } 2i < n, \\ \mathcal{V}(n - 2i) \oplus \mathcal{V}(2n - 2i), & \text{if } 2i > n, \\ \mathcal{V}(0), & \text{if } i = 1 \text{ and } n = 2, \end{cases}$$

for the spaces $\mathcal{V}(j)$ introduced in Theorem 2.8.

Then the span of words in J_n^i and $(J_n^T)^{n-i}$ of length $n - 2$ is a proper subspace of the space $\mathcal{W}(n, i)$.

Proof. For the sake of simplicity let us denote $A_n = J_n^i, B_n = (J_n^T)^{n-i}$.

Take an arbitrary word in A_n and B_n of length $n - 2$:

$$A_n^{p_1} B_n^{t_1} A_n^{p_2} B_n^{t_2} \dots A_n^{p_s} B_n^{t_s}, \tag{6}$$

where p_1, t_s are nonnegative and t_l, $l = 1, 2, \dots, s - 1$, p_l, $l = 2, \dots, s$ are positive integers. Set $p = \sum_{l=1}^s p_l, t = \sum_{l=1}^s t_l$. Then $p + t = n - 2$ and Lemma 2.7 implies that a word of the form (6) can have nonzero entries only on $((i - t)n - 2i)$-diagonal (since $ip - (n - i)t = i(n - 2 - t) - (n - i)t = (i - t)n - 2i$).

If $2i = n$, then we need to have $i = 1$ and $n = 2$, and we obtain the word of length zero I_2.

For $2i < n$ we have $|(i - t)n - 2i| \leq n - 1$ if and only if either $t = i$ or $t = i - 1$. That is nontrivial words of length $n - 2$ have nonzero elements either on $(-2i)$- or $(n - 2i)$-diagonal, thus are contained in $\mathcal{W}(n, i)$.

For $2i > n$ we have $|(i - t)n - 2i| \leq n - 1$ if and only if either $t = i - 2$ or $t = i - 1$. That is nontrivial words of length $n - 2$ have elements either on $(n - 2i)$- or $(2n - 2i)$-diagonal, thus are also contained in $\mathcal{W}(n, i)$.

Consequently, the span of all words in A_n and B_n of length $n - 2$ is contained in $\mathcal{W}(n, i)$.

Now we are left to prove that all words in A_n and B_n of length $n - 2$ span a subspace in $\mathcal{W}(n, i)$, which does not coincide with the whole space $\mathcal{W}(n, i)$.

We prove this statement using the induction on n.

Induction base: If $n = 2$, the only possible value for i is 1. Since $n - 2 = 0$, the only word of length $n - 2$ is thus I_2, and since $\dim \mathcal{V}(0) = n = 2$, the base of induction follows.

Induction step: Let $n \geq 3$ and assume that for all $n' < n$ the statement holds.

I. If $i = 1$ and $n \geq 3$, the result follows from the proof of [9, Theorem 2, case $p = n - 1$] for arbitrary n.

Consider $i \geq 2$. Let us write n as $n = qi + r$, where $1 \leq r < i$.

II. Assume that $2 \leq i < \frac{n}{2}$.

In this case we have $B_n^2 = 0$ and $A_n^{q+1} = 0$, and a nonzero word (6) is of the form

$$A_n^{p_1} B_n A_n^{p_2} B_n \ldots A_n^{p_s},$$

where $p_l \leq q$ for all $l = 1, \ldots, s$ and $s = i$ or $s = i + 1$.

Notice that all elements of a product $B_n A_n^j$, $j = 0, \ldots, q$, are equal to zero, except those in the last i rows. The latter have the following form:

$$\begin{pmatrix} I_i & O_i & \ldots & O_i & O_{i,r} \end{pmatrix},$$
$$\begin{pmatrix} O_i & I_i & \ldots & O_i & O_{i,r} \end{pmatrix},$$
$$\vdots$$
$$\begin{pmatrix} O_i & O_i & \ldots & I_i & O_{i,r} \end{pmatrix},$$
$$\begin{pmatrix} O_i & O_i & \ldots & O_i & I_{i,r} \end{pmatrix},$$

where $O_{i,r}$ and $I_{i,r}$ denote $i \times r$ submatrices of O_i and I_i, respectively, formed by first r columns. Now we see that for $j < q - 1$, the last i columns of $B_n A_n^j$ are zero, and this gives us

$$B_n A_n^j B_n = 0, \ j = 1, \ldots, q - 2.$$

Therefore a word V of length $n - 2$, which is not trivially zero and has its nonzero elements located on $(n - 2i)$-diagonal, has the form

$$A_n^{p_1} B_n A_n^{q-\varepsilon_1} B_n \cdots A_n^{q-\varepsilon_{i-2}} B_n A_n^{p_i},$$

$\varepsilon_j \in \{0, 1\}$, $j = 1, \ldots, i - 2$.

Consider the projections of $B_n A_n^{q-1}$ and $B_n A_n^q$ on the lower right $i \times i$ corner. They are $(J_i^T)^r$ and J_i^{i-r}, respectively. Therefore a subword

$$B_n A_n^{q-\varepsilon_1} B_n \cdots B_n A_n^{q-\varepsilon_{i-2}}$$

of V corresponds to a word of length $i-2$ in $(J_i^T)^r$ and J_i^{i-r}. By construction $\gcd(i, r) = 1$, hence by induction hypothesis the span of all words of length $i - 2$ in $(J_i^T)^r$ and J_i^{i-r} is a proper subspace $\mathcal{W}$ of the space $\mathcal{W}(i, i - r) \subset$

$M_i(\mathbb{C})$. Take the words $U_1, \ldots, U_s, U_{s+1}, \ldots, U_m$, which form a basis of the subspace $\mathcal{W}$, where $m < \dim \mathcal{W}(i, i - r)$, $U_1, \ldots, U_s$ are lower-triangular, $U_{s+1}, \ldots, U_m$ are upper-triangular. They have pre-images

$$V_1 = B_n A_n^{q - \varepsilon_{1,1}} B_n \cdots A_n^{q - \varepsilon_{i-2,1}}, \ldots, V_m = B_n A_n^{q - \varepsilon_{1,m}} B_n \cdots A_n^{q - \varepsilon_{i-2,m}},$$

$\varepsilon_{a,b} \in \{0, 1\}$ for all $a = 1, \ldots, i-2$, $b = 1, \ldots, m$. The words $V_1 B_n, \ldots, V_m B_n$ span a space equal to $\mathcal{W}$ located in the left lower $i \times i$ corner of $n \times n$ matrices. Finally we consider all words $A_n^{p_1} V_j B_n A_n^{p_i}$ of length $n - 2$.

1. Suppose $2r < i$. The intersection of $\mathcal{V}(n - 2i)$ with the space generated by $i \times 2i$ matrices formed by rows $r+1, \ldots, r+i$ and columns $(q-2)i+1, \ldots, qi$ is isomorphic to $\mathcal{W}(i, i - r)$. To get nonzero entries in this submatrix we are able to take the words $A_n^{q-1} V_h B_n A_n^{q-2}$, $h \geq s + 1$ and $A_n^{q-1} V_g B_n A_n^{q-1}$, $1 \leq g \leq s$. By construction

$$\dim \langle A_n^{q-1} V_h B_n A_n^{q-2}, \ A_n^{q-1} V_g B_n A_n^{q-1}, s + 1 \leq h \leq m, 1 \leq g \leq s \rangle = m <$$
$$< \dim \mathcal{W}(i, i - r),$$

thus the words in A_n and B_n of length $n - 2$ do not span $\mathcal{W}(n, i)$.

2. Suppose $2r = i$, that is $r = 1$, $i = 2$. The only nontrivial word of length $n - 2$ with 1 letter B_n which lies in the 2-dimensional intersection of $\mathcal{V}(n - 2i)$ with the space generated by 2×2 matrices formed by rows $r + 1, r + 2$ and columns $2q - 1, 2q$ is $A_n^{q-1} B_n A_n^{q-1}$. Thus the words in A_n and B_n of length $n - 2$ do not span $\mathcal{W}(n, 2)$.

3. Suppose $2r > i$. The intersection of $\mathcal{V}(n - 2i)$ with the space generated by $i \times (i+r)$ matrices formed by rows $r+1, \ldots, r+i$ and last $i+r$ columns is isomorphic to $\mathcal{W}(i, i-r)$, since in $M_i(\mathbb{C})$ we have $\dim \mathcal{V}(-2(i-r)) = 2r - i < r$. To get to this submatrix we are able to take the words $A_n^{q-1} V_h B_n A_n^{q-1}$, $h \geq s + 1$ and $A_n^{q-1} V_g B_n A_n^q$, $1 \leq g \leq s$. By construction

$$\dim \langle A_n^{q-1} V_h B_n A_n^{q-1}, \ A_n^{q-1} V_g B_n A_n^q, s + 1 \leq h \leq m, 1 \leq g \leq s \rangle = m <$$
$$< \dim \mathcal{W}(i, i - r),$$

thus the words in A_n and B_n of length $n - 2$ do not span $\mathcal{W}(n, i)$.

III. If $i > \frac{n}{2}$, then we apply the arguments from I or II to matrices A_n^T and B_n^T, to show that their products of length $n - 2$ span a proper subspace $\mathcal{W}$ in $\mathcal{W}(n, n - i)$. Thus we take transposes again, and see that A_n and B_n generate a proper subspace $(\mathcal{W})^T$ in $(\mathcal{W}(n, n-i))^T = \mathcal{W}(n, i)$ and the result follows.

$\square$

Theorem 2.11. *Let $i < n$ be positive integers, $\gcd(n, i) = d$. Then for matrices from $M_n(\mathbb{C})$ we have*

$$l(\{J_n^i, (J_n^T)^{n-i}\}) = l_0(\{J_n^i, (J_n^T)^{n-i}\}) = 2\frac{n}{d} - 2.$$

Proof. It follows from the proof of Theorem 2.9, that if $\gcd(n, i) = d > 1$, then $l(\{J_n^i, (J_n^T)^{n-i}\}) = l(\{J_{\frac{n}{d}}^{\frac{i}{d}}, (J_{\frac{n}{d}}^T)^{\frac{n-i}{d}}\})$, therefore it is sufficient to consider the case when i and n are co-prime.

Trivially we have $l(\{J_n^i, (J_n^T)^{n-i}\}) = l(\{J_n^i, J_n^i + (J_n^T)^{n-i}\})$. Since $J_n^i + (J_n^T)^{n-i}$ is a non-derogatory matrix, which has n distinct roots of unity as its eigenvalues, then the upper bound $l(\{J_n^i, J_n^i + (J_n^T)^{n-i}\}) \leq 2n - 2$ follows from [12, Theorem 4.1. (b)].

Now we prove the lower bound $l(\{J_n^i, (J_n^T)^{n-i}\}) \geq 2n - 2$.

Without loss of generality we assume now that $i < \frac{n}{2}$, since

$$l(\{J_n^i, (J_n^T)^{n-i}\}) = l(\{(J_n^T)^i, J_n^{n-i}\}).$$

Arguing similarly as when calculating the period of the sequence (4) in Theorem 2.9, and applying Lemma 2.7 we calculate that a word with p letters J_n^i and t letters $(J_n^T)^{n-i}$ has its nonzero entries on $-2i$- or $(n - 2i)$-diagonal, i.e. belongs to the space $\mathcal{W}(n, i)$, if and only if $p + t + 2 = \alpha n$ for positive integers α. Lemma 2.10 yields that considering only $\alpha = 1$ is not enough to generate $\mathcal{W}(n, i)$, thus it is necessary to consider $\alpha = 2$ as well. Consequently, there exists a word of length $2n - 2$ which is not contained in $\mathcal{L}_{2n-3}(\{J_n^i, (J_n^T)^{n-i}\})$ and $l(\{J_n^i, (J_n^T)^{n-i}\}) \geq 2n - 2$.

Since $J_n^i + (J_n^T)^{n-i}$ is an invertible matrix, the equation $l(\{J_n^i, (J_n^T)^{n-i}\}) = l_0(\{J_n^i, (J_n^T)^{n-i}\})$ follows from Proposition 2.3. $\square$

One of Hans Schneider's first papers *A pair of matrices with property P* [14] deals with conditions which guarantee the simultaneous triangularizability of a set of matrices, and he continued to contribute to, and influence research in matrix theory from the standpoint of the theory of finite dimensional algebras.

References

References

[1] D. Constantine and M. Darnall. Lengths of finite dimensional representations of PBW algebras. *Linear Algebra Appl.*, 395:175–181, 2005.

[2] Fergus J. Gaines, Thomas J. Laffey, and Helene M. Shapiro. Pairs of matrices with quadratic minimal polynomials. *Linear Algebra Appl.*, 52/53:289–292, 1983.

[3] A. E. Guterman and O. V. Markova. Commutative matrix subalgebras and length function. *Linear Algebra Appl.*, 430(7):1790–1805, 2009.

[4] A. E. Guterman, O. V. Markova, and S.D. Sochnev. The algebra of semimagic matrices and its length. *Zap. Nauch. Sem. POMI*, 419:52–76, 2013. Translated in J. Math. Sci., New York, 2014, 199:4, 400–413.

[5] Thomas J. Laffey. Algebras generated by two idempotents. *Linear Algebra Appl.*, 37:45–53, 1981.

[6] M. S. Lambrou and W. E. Longstaff. On the lengths of pairs of complex matrices of size six. *Bull. Aust. Math. Soc.*, 80(2):177–201, 2009.

[7] W. E. Longstaff. Burnside's theorem: irreducible pairs of transformations. *Linear Algebra Appl.*, 382:247–269, 2004.

[8] W. E. Longstaff, A. C. Niemeyer, and Oreste Panaia. On the lengths of pairs of complex matrices of size at most five. *Bull. Austral. Math. Soc.*, 73(3):461–472, 2006.

[9] W. E. Longstaff and Peter Rosenthal. On the lengths of irreducible pairs of complex matrices. *Proc. Amer. Math. Soc.*, 139(11):3769–3777, 2011.

[10] O. V. Markova. Computation of the lengths of matrix subalgebras of a special form. *Fundam. Prikl. Mat.*, 13(4):165–197, 2007. Translated in J. Math. Sci., New York, 2008, 155:6, 908–931.

[11] O. V. Markova. The length function and matrix algebras. *Fundam. Prikl. Mat.*, 17(6):65–173, 2011/12. Translated in J. Math. Sci., New York, 2013, 193:5, 687–768.

[12] Christopher J. Pappacena. An upper bound for the length of a finite-dimensional algebra. *J. Algebra*, 197(2):535–545, 1997.

[13] Azaria Paz. An application of the Cayley-Hamilton theorem to matrix polynomials in several variables. *Linear and Multilinear Algebra*, 15(2):161–170, 1984.

[14] Hans Schneider. A pair of matrices with property P. *Amer. Math. Monthly*, 62:247–249, 1955.